

Chapter 3

Addressing DNA Origami with Polyamide-Biotin Conjugates

Abstract

DNA origami is a unique DNA architecture that can be used to create arbitrary two-dimensional shapes on the nanoscale. As with all DNA nanoarchitectures, the ability to create functional molecular assemblies from this DNA template will be dependent on the ability to recruit active molecules to the DNA surface in a bottom-up approach to self assembly. The ability to target these unique DNA nanostructures with polyamide-biotin conjugates and recruit streptavidin to their surface was examined using atomic force microscopy. Evidence for recruitment at predicted binding sites and an outline for future work is presented.

3.1 Introduction

The ability to create DNA nanostructures containing precise spacing's and shapes is an important requirement for bottom-up self assembly.¹⁻⁴ In this approach, DNA templates are used for the organization of secondary molecular components. To this end, DNA origami has been shown to be an ideal method for the creation of arbitrary 2-dimensional shapes.⁵ As originally demonstrated by Paul Rothemund, DNA origami entails the folding of a long, single-stranded DNA scaffold strand into a variety of shapes by the addition of a large number (typically >200) of short oligonucleotide staple strands. The scaffold strand used is the 7,249 nt long M13mp18 viral DNA strand. The staple strands are typically 32 nt long and can be used without purification. A variety of different shapes have been generated using this method including squares, rectangles, five-point stars, smiley faces, and even a map of China.^{5, 6} The origami itself can be decorated with DNA dumbbells by modifying the staple strands, allowing words such as "DNA" and even pictures to be drawn on the surface of the origami.⁵ DNA origami can also be modified to display single stranded DNA sticky ends on the edges, as well as on the surface.^{7, 8}

For DNA origami to act as a template for increasingly complex assemblies, a method of recruiting active components to the DNA structure is needed. Given the previous success in using polyamides to recruit streptavidin to DX arrays described in the previous chapter,^{9, 10} and the structural similarity between origami and the DX molecule, the ability of polyamide conjugates to functionalize DNA origami was next investigated.

3.2 Results and Discussion

Analysis of Binding Sites on DNA Origami Smiley Face

For this study, the DNA origami smiley face was used as originally described by Rothemund.⁵ Although structurally related, DX arrays and DNA origami have several significant differences. The periodic nature of DX arrays results in the presentation of a relatively small number of DNA sequences. In contrast, assuming that the full scaffold strand is used, over 7,000 base pairs are present in a DNA origami structure, virtually ensuring that multiple binding sites will be present for a standard pyrrole-imidazole eight-ring polyamide that has specificity for 6 base pairs. Similarly, the use of the viral DNA scaffold strand also makes modification of the origami sequence difficult. When designing a DX tile one can ensure that a specific sequence occurs only a single time throughout the array. The ease of synthesis of short oligonucleotides allows one to easily manipulate the sequences in the array to meet this requirement. However, extensive




Polyamide	Sequence	# Sites in Full Sequence	Accessibility
 1	5'-WGGWCW-3'	24	7 sites - Optimal 17 sites - at nick / junction / edge
 2	5'-WTWCGW-3'	22	
 3	5'-WGWGCW-3'	14	7 sites - Optimal 7 sites - at nick / junction / edge

Figure 3.1 Analysis of potential polyamide binding sites on a DNA origami smiley face. Polyamides **1 – 3** are shown along with their target sequences. The number of match sites found throughout the structure is shown, as well as how many of the sites are presumed to be accessible. Sites are deemed optimal unless any of the base pairs are located at a junction, have a nick, or occur at the edge of the structure.

modification of the origami scaffold strand would be challenging as its length precludes the use of standard solid phase DNA synthesis to generate a new strand.

An analysis of the smiley face origami was done to determine how many potential binding sites were present for three previously characterized polyamide-biotin conjugates that were used in targeting DX-arrays.¹⁰ As shown, polyamide **1** is specific for the sequence 5'-WGGWCW-3',¹¹ polyamide **2** is specific for 5'-WTWCGW-3',¹² and polyamide **3** targets 5'-WGWGCW-3'.¹³ Due to the lack of specificity observed for polyamide **2** in previous studies,^{10, 14} only polyamides **1** and **3** were chosen for further

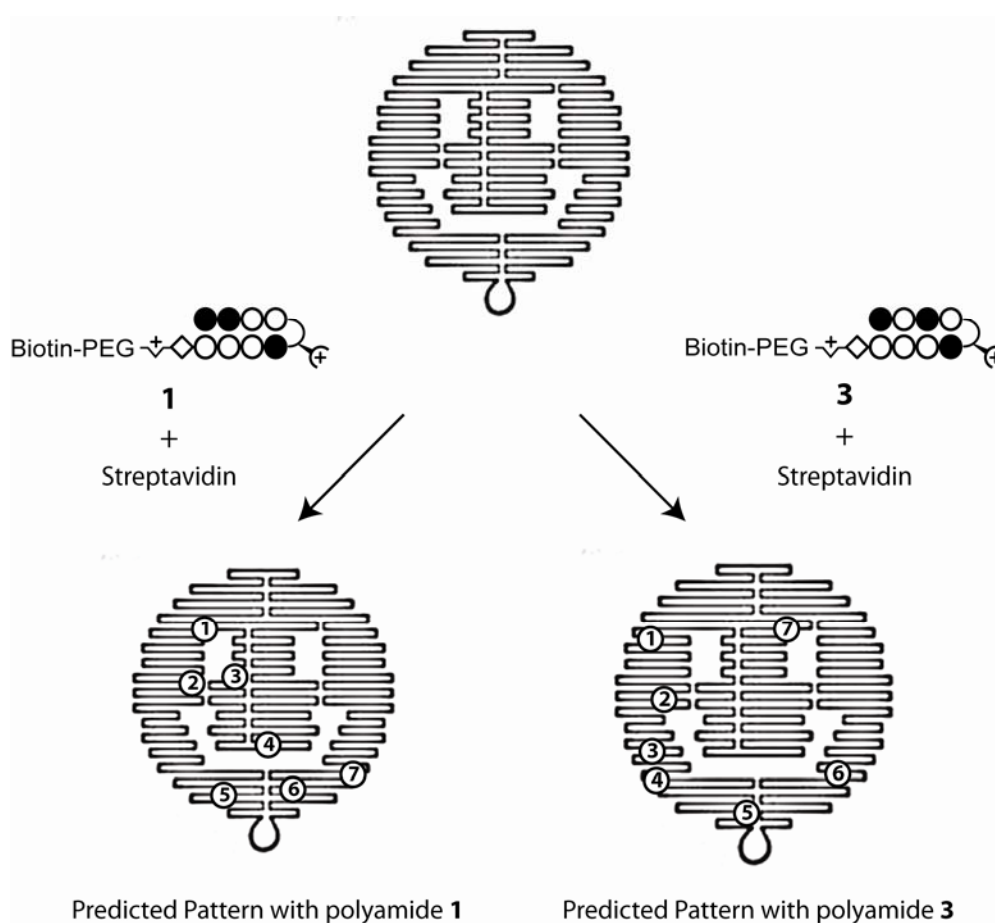


Figure 3.2 Predicted binding sites on DNA origami for polyamides **1** and **3**. The 7 optimal binding sites for each polyamide are indicated.

experiments. The total number of binding sites for their target sequences was determined as well as whether the sites were likely to be accessible and capable of being bound by the polyamide. Previous work had shown that polyamides could not bind to the crossover junction in a DX molecule.⁹ As a result binding sites that occurred at the various nicks that occur between staple strands, at any of the crossover points, or at the edges were deemed inaccessible. All other sites were considered to be optimal for binding. The location of each of the optimal binding sites was mapped to the smiley face. In both cases, there were 7 optimal sites that are predicted to create a unique pattern of observable “bumps” when the polyamide-biotin conjugates are incubated with streptavidin and the origami. The predicted patterns are shown in Figure 3.2.

Initial AFM Experiments with DNA Origami

The initial AFM experiments reproduced the creation of smiley face origami structures. Well-formed structures in agreement with the previous work were observed. The quality of structures was observed to decrease over time, so experiments were done within 24 hours of annealing. A control with streptavidin added to the origami was done to ensure that non-specific binding between the streptavidin and the origami did not occur. Next, polyamides **1** and **3** (100 nM) were separately incubated with the DNA origami (1.8 nM) and streptavidin (200 nM). A large sample of images were taken in order to determine how well the observed patterns fit with those that were predicted. In most images the smileys were observed clustered together. This is likely due to base-stacking at the ends of the origami as was observed in the original study, and could

potentially be eliminated by the addition of 4T loops at the ends of the helices. In the case of polyamide **3**, an extremely large cluster of smileys was observed. As shown in Figure

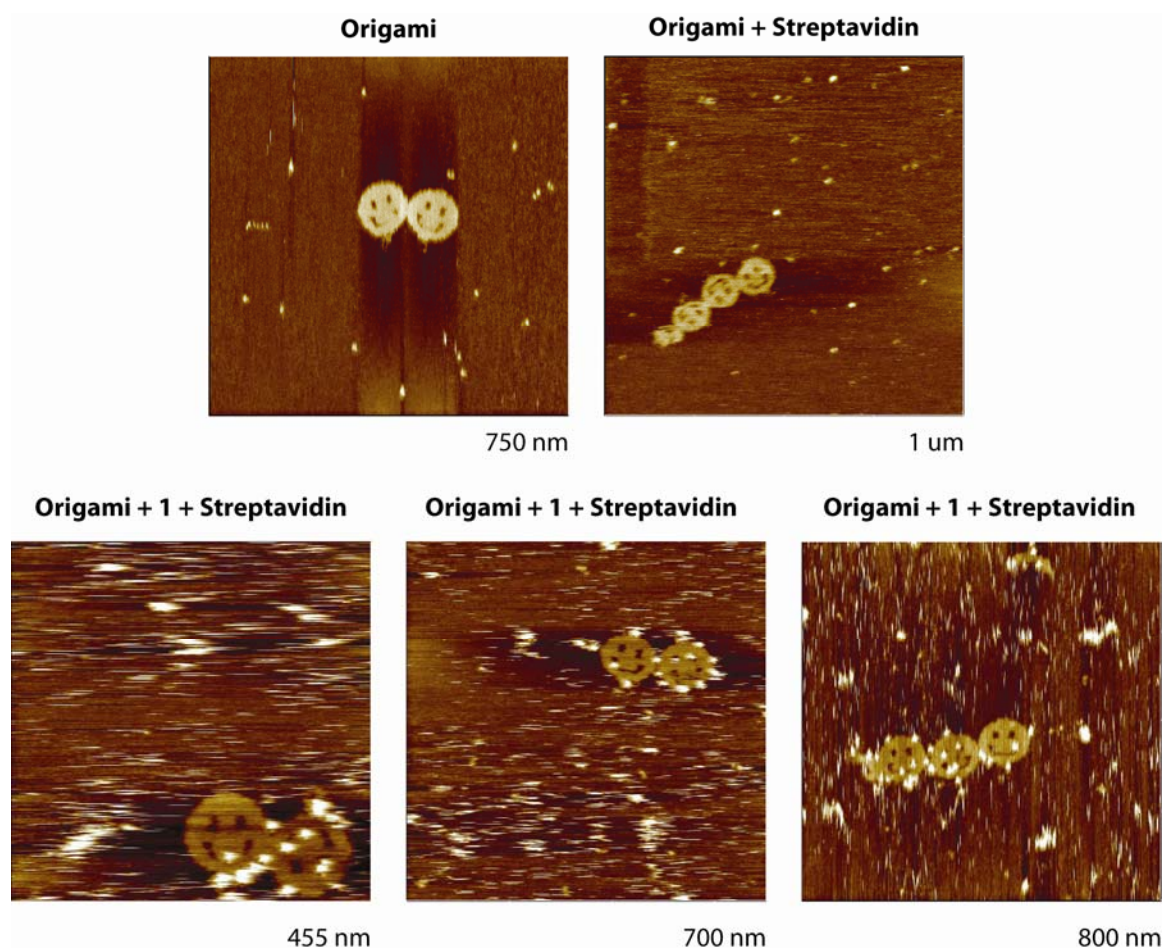


Figure 3.3 AFM Images of DNA origami with and without polyamide **1**. The dimensions of the image are shown. The top row contains the origami only and origami + streptavidin controls. The bottom row contains origami + **1** + streptavidin.

3.4, thirteen separate $1\ \mu\text{m} \times 1\ \mu\text{m}$ AFM scans were used to collect images of the entire cluster. The images were then merged, giving a set of 235 distinct DNA origami Smileys. Of these, 66 were unlabelled, 95 contained 1 bump, 58 had 2 bumps, 12 had 3 bumps, and 4 had 4 or more bumps. 72% contained at least 1 bump, and on average 1.13 bumps per structure were observed.

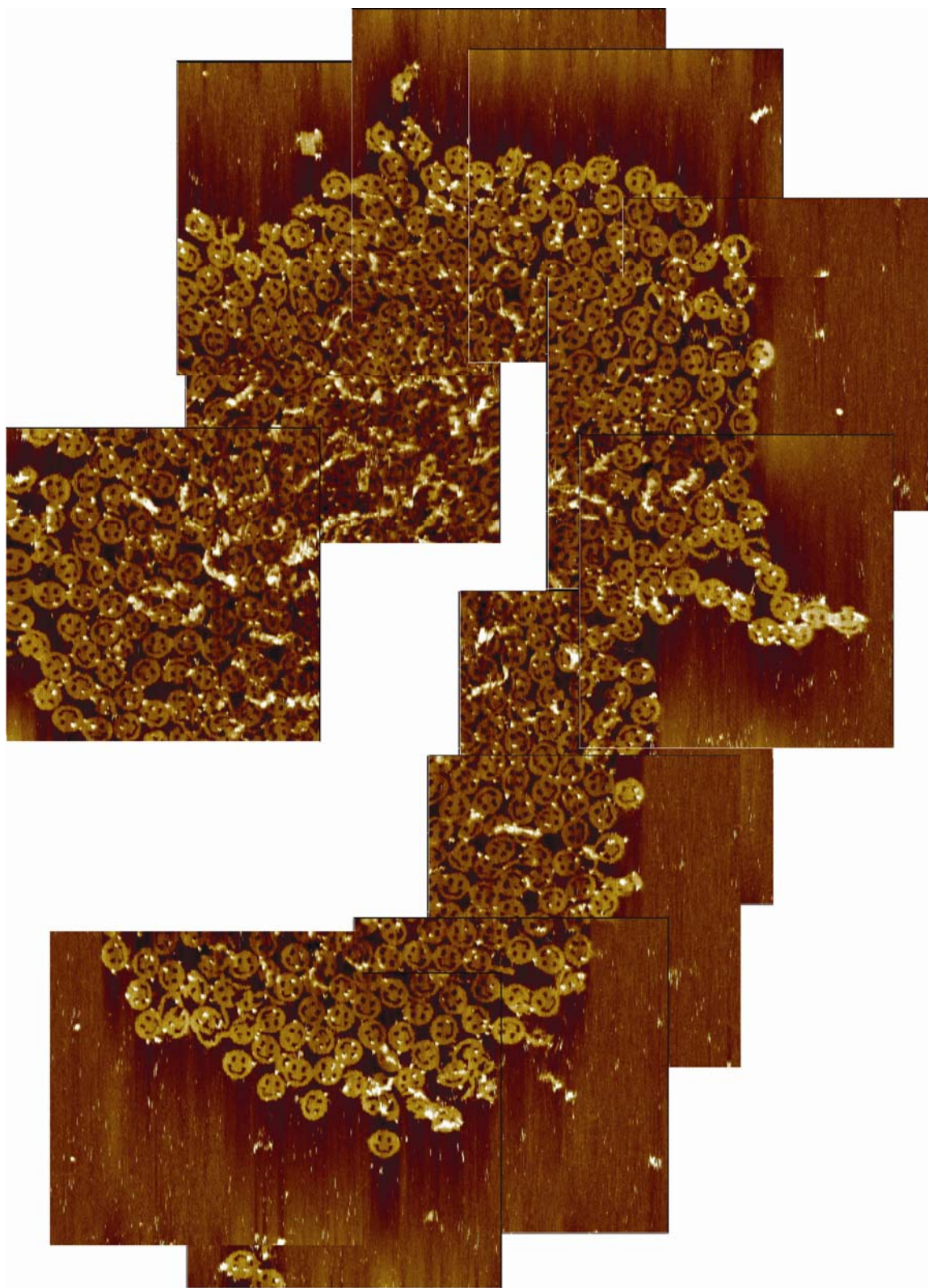


Figure 3.4 AFM Images of DNA origami with polyamide **3**. Thirteen individual 1 μm x 1 μm scans were performed and the images overlapped. 235 distinct structures were observed and used to determine binding locations as well as estimate the efficiency of labeling.

Analysis of AFM Experiments

Using all of the collected images, the individual patterns observed were matched with the predicted patterns for each set of polyamides to determine if there was a correlation between the observed and predicted protein recruitment. As shown in Figures 3.6 and 3.7, individual binding events were observed that appeared to occur at all of the predicted locations. However, analysis of the results was greatly hampered due to the macromolecular symmetry of the smiley face DNA structure. With a symmetric structure and no inherent topological marker, it becomes impossible to determine whether any given observed structure is in a face-up or face-down orientation. The symmetry of the

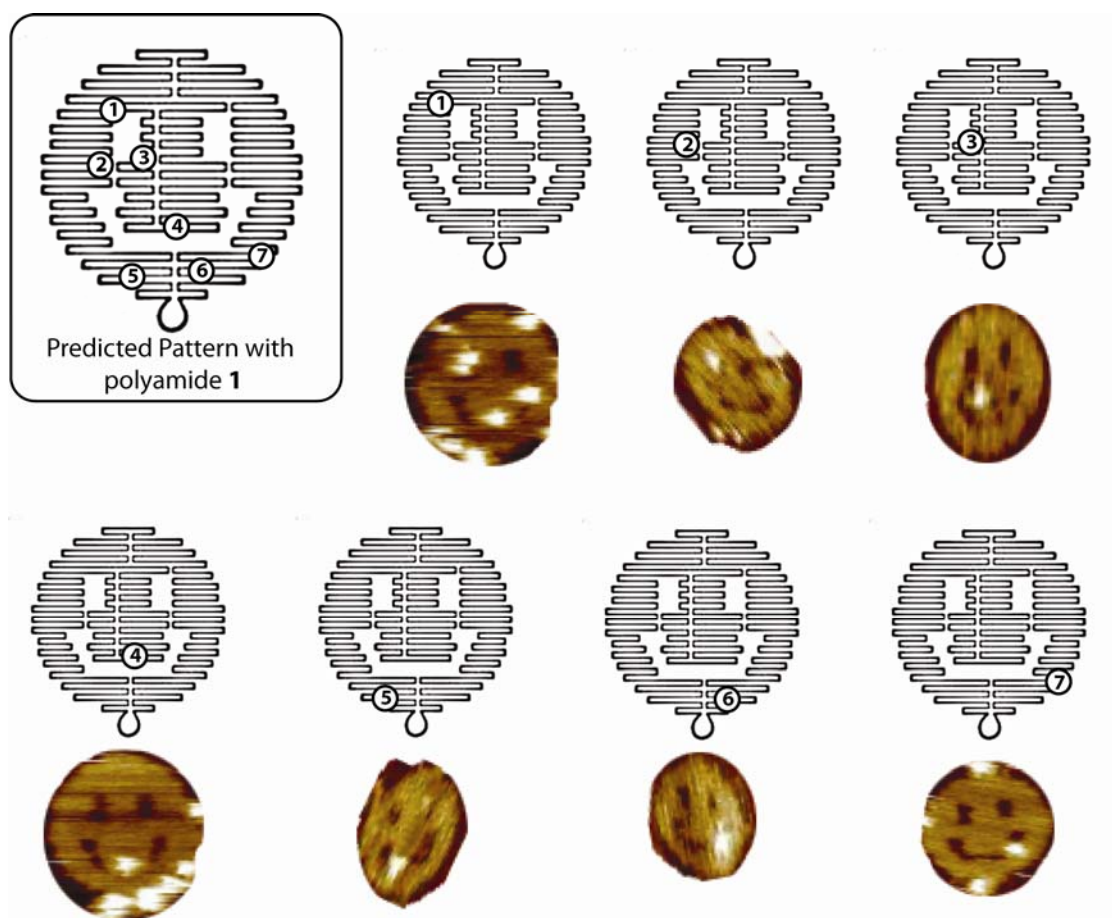


Figure 3.5 Comparison of predicted and observed binding patterns for polyamide 1. The locations of the seven optimal binding sites are shown along with images that have binding in the approximate location.

structure makes it especially difficult to distinguish between sites that are effectively mirror images of each other such as sites 5 and 6 with polyamide **1** or 4 and 6 with polyamide **3**. Exact determination of the binding sites was additionally complicated by the relatively low level of labeling observed, as attempting to decipher the orientation based upon the location of only one or two bumps was difficult. As seen in the numerous images in Figures 3.3 – 3.7, the number of binding events per structure was quite limited and

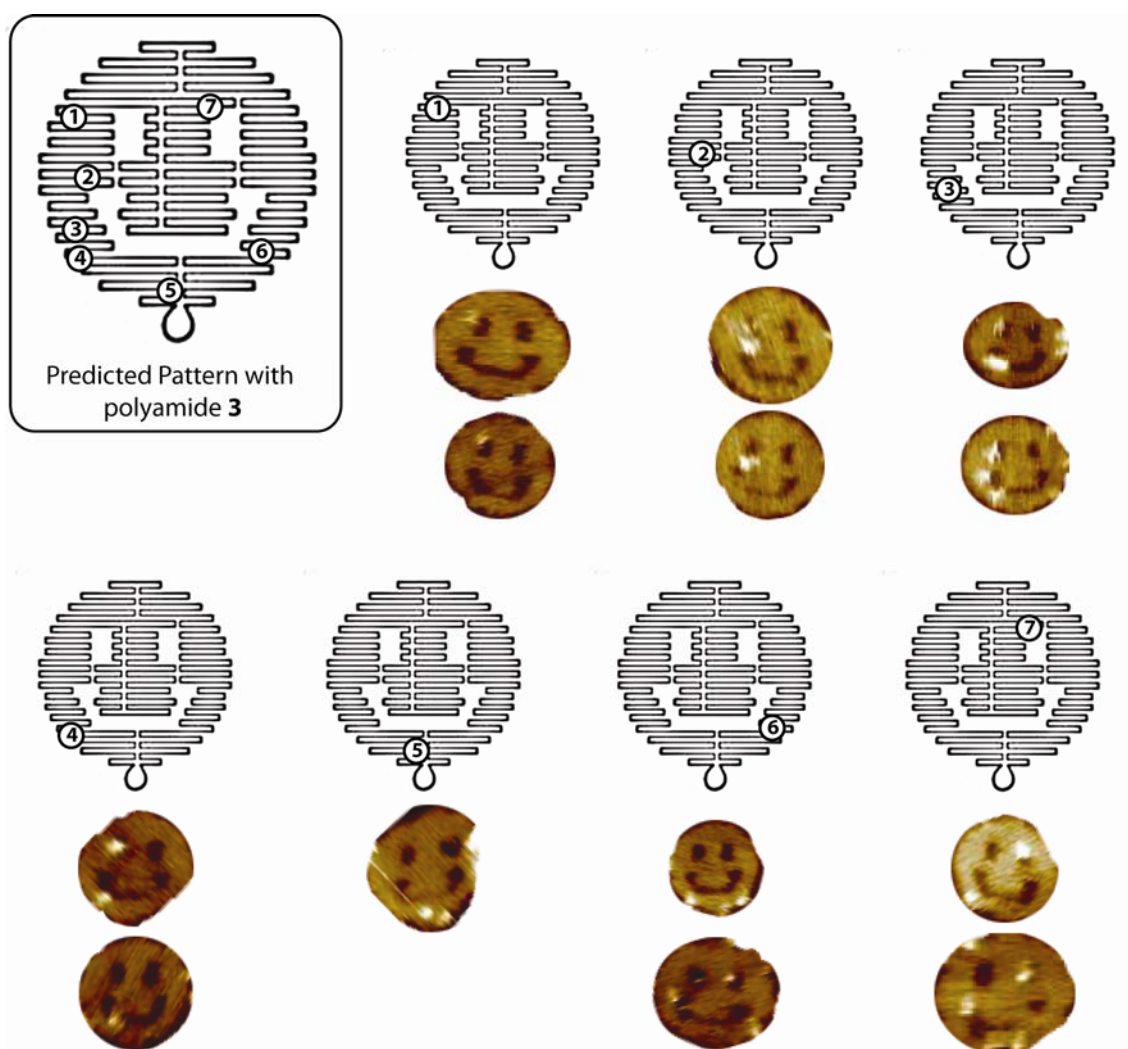


Figure 3.6 Comparison of predicted and observed binding patterns for polyamide **3**. The locations of the seven optimal binding sites are shown along with images that have binding in the approximate location.

in no case was binding at all 7 sites observed for an individual structure. Finally, the spatial resolution made it difficult to unambiguously differentiate between binding sites that were in close proximity such as sites 3 and 4 with polyamide **3**.

Asymmetric DNA Origami

As discussed in the previous section, the symmetry of the DNA origami smiley face greatly complicated the ensuing analysis of the data. As such, several strategies to overcome this difficulty were examined. Although asymmetric origami structures are known,⁶ the complete redesign and synthesis of a new asymmetric smiley face was deemed unpractical due to the time and cost involved in the de novo synthesis of a new origami structure. A second potential strategy was to take advantage of the fact that although the overall macrostructure of the smiley face origami is symmetric, the path which the scaffold strand traces and the microstructure is not. High quality AFM scans could potentially show the underlying microstructure of the origami.⁵ However, this is a highly time consuming process and would not be suitable for the analysis of a large number of structures. The most practical approach and the one which was pursued is to incorporate DNA hairpins into the origami structure. Two of the standard staple strands were replaced with staples containing DNA dumbbells. These dumbbells were designed to function as a constant topographic marker which could be used to unambiguously determine the orientation of each smiley face. The modified staples were located in the top right of the smiley face as this location has no predicted binding sites for either polyamide. The two modified staple strands were s3m4e and s5m4e (see ref. 5, Supporting Information for the exact location) which both span the 4th and 5th helix from

the top of the structure and are located above and to the right of the smiley's right "eye". These staple strands were modified by the insertion of the sequence 5'-TCCTCTTTTGAGGAACAAGTTTCTTGT-3' between the 16th and 17th nucleotides as was previously done to decorate origami with DNA dumbbells. The dumbbell structure is thought to be superior to simple DNA hairpins for this application as it avoids unwanted intermolecular dimerization.⁵

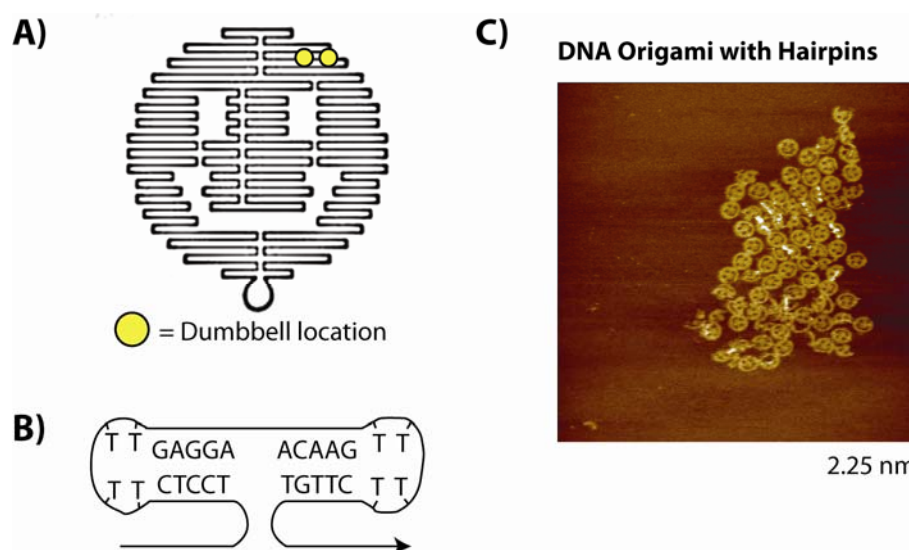


Figure 3.7 Breaking the symmetry of the DNA smiley. A) The location of the two DNA dumbbells is shown. B) The sequence and structure of the DNA dumbbells. C) AFM image of the modified smiley.

As shown in Figure 3.8C the DNA dumbbells were not visible in the modified origami. One possible explanation for this is that a larger number of dumbbells are necessary for visualization. In the case of modified DX arrays or the previously modified origami, the large number of DNA barbells in close proximity with each other may act to reinforce and rigidify each other due to the steric crowding. In the case of only two DNA dumbbells, one expects that they might be "floppier" and less visible using AFM. It should be noted that in modifying DNA origami rectangles with dumbbells to create an orientation point and to encode separate structures, the Yan group used clusters of 4

dumbbells. Future work would likely benefit from the incorporation of 2 to 4 additional DNA dumbbells in the vicinity of the original 2 to make an easily distinguishable topographical point for orienting the origami structures.

3.3 Conclusions

The use of polyamides to attach protein to DNA origami has been initially explored on a DNA smiley face. However several challenges remain for this approach to have practical uses. First, successful modification of the origami with DNA dumbbells in order to break the symmetry of the macrostructure is necessary in order to unambiguously interpret the binding patterns of the polyamides. As this has been previously demonstrated, it is not anticipated that this would be an insurmountable challenge. The greater difficulties with this approach are the specificity and labelling efficiency of polyamides. It is clear from the analysis shown in Figure 1 that standard eight-ring polyamides do not have enough intrinsic specificity to unambiguously target a single sequence in a large DNA origami structure. One potential solution is to use polyamides containing higher imidazole content, as these would be more specific than the polyamides used in this study. Two such polyamides are shown in Figure 9, which would target the sequences 5'-WGGGGW-3' and 5'-WGGGCW-3' respectively. It is worth noting that the analogous polyamides without biotin attached have been shown to target these sequences with high affinity. As shown, these polyamides are expected to have only

Polyamide	Sequence	# Sites in Full Sequence	Accessibility
	5'-WGGGGW-3'	3	2 sites - Optimal
	5'-WGGGCW-3'	9	1 site - Optimal

Figure 3.8 Polyamides proposed to have improved specificity for DNA sequences in the DNA smiley.

2 and 1 optimal binding sites respectively in the DNA smiley. Similarly, another potential solution would be to use extended polyamide motifs which are capable of binding to more than 6 base pairs.

The other major difficulty facing this approach is the low level of labelling observed by the polyamide-biotin conjugates. As previously discussed, examination of 235 unique smileys found on average just over 1 binding event per structure. It is surprising given the presence of multiple match binding sites, as well as the presence of numerous weak mismatch binding sites that so little binding is observed. Also of concern is the highly variant nature of binding from structure to structure. Although it is possible this is in part due to the inherent damage to the structures from AFM imaging, the relative stability of the structures when scanned multiple times indicates that AFM tip damage does not fully account for the observed results.

It is worth noting that in examining the binding of DNA oligomers to ssDNA overhangs on the face of DNA origami rectangles, the Yan lab found that different locations on the surface of the origami modified with the same sequence had markedly different labelling efficiencies.⁸ Positions near the edge had much higher efficiency binding than those located towards the center of the structure. These results clearly show

that every location on the DNA origami surface is not equal in terms of accessibility. Given the drastic differences observed for the binding of DNA oligomers, it is likely that proteins being recruited to the surface of the origami would be similarly affected. The presence of several large “holes” in the surface of the DNA smiley face may also play a role in determining which sites are most accessible for binding.

Addressing the challenge of highly efficient labelling will likely require further understanding of how binding occurs on the surface of these DNA structures. It is also anticipated that polyamides with higher affinities and specificities for DNA may be beneficial for labelling DNA nanostructures with improved efficiency.

3.4 Experimental Details:

Abbreviations. Trishydroxymethylaminomethane (TRIS), ethylenediaminetetraacetic acid (EDTA).

Materials. Boc- β -Ala-PAM resin was purchased from Peptides International. Trifluoroacetic acid (TFA) was purchased from Halocarbon. Methylene Chloride (DCM) was obtained from Fisher Scientific and *N,N*-dimethylformamide (DMF) from EMD. EZ-Link TFP-PEO₃-Biotin was purchased from Pierce. Streptavidin was purchased from Rockland. Ultra Pure TRIS was purchased from ICN. Magnesium Acetate 4-hydrate was obtained from J.T. Baker. Water (18 M Ω) was purified using a aquaMAX-Ultra water purification system. Biological experiments were performed using Ultrapure Water (DNase/RNase free) purchased from USB. The pH of buffers was adjusted using a

Thermo Orion 310 PerpHect Meter. All buffers were sterilized by filtration through a Nalgene 0.2 μm cellulose nitrate filtration.

Polyamide Synthesis. Polyamide monomers were prepared as described previously.¹⁵ Synthesis was performed using established protocols and all polyamides were characterized by MALDI-TOF and analytical HPLC. The synthesis of polyamides **1** and **3** is described in section 2.4 and has been published.¹⁰

DNA Oligomers. DNA oligomers were purchased from IDT Technologies in a 96 well format. The oligomers were ordered on 100 nmol scale, unpurified, and at 150 μM concentration in water. The sequences of the 243 staple strands are described in Rothmund's original paper.⁵ The M13mp18 ssDNA scaffold strand was purchased from New England Biolabs.

Annealing DNA Origami. The DNA origami was annealed as follows. First, 5 μL of each of the staple strands was mixed to create a 10X stock solution (620 nM). Fresh origami samples were annealed 24 hours before AFM experiments as follows. 10 μL of the 10X stock of staple strands (62 nM) was mixed with 1 μL of the scaffold strand (1.3 nM) in TAEMg Buffer (40 mM Tris-HCl (pH 8.0), 20 mM acetic acid, 1 mM EDTA, 12.5 mM magnesium acetate) in a final volume of 100 μL . The sample was placed on a float in a styrofoam box filled with 1 L of water at 95°C and allowed to cool to room temperature slowly overnight. The DNA dumbbell modified smileys were made in the same fashion but the modified oligomers were used instead of the original staple strands.

AFM Sample Preparation. DNA origami was incubated with 100 nM of polyamide **1** or **3** and 200 nM Streptavidin prior to imaging. 2.5 μ L of sample was spotted on mica and imaged under TAEMg buffer.

3.5 References

1. Seeman, N. C.; Lukeman, P. S., Nucleic acid nanostructures: bottom-up control of geometry on the nanoscale *Rep. Prog. Phys.* **2005**, 68, 237-270.
2. Lin, C. X.; Liu, Y.; Rinker, S.; Yan, H., DNA tile based self-assembly: Building complex nanoarchitectures *ChemPhysChem* **2006**, 7, 1641-1647.
3. Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F., Assembling materials with DNA as the guide *Science* **2008**, 321, 1795-1799.
4. Seeman, N. C., DNA in a material world *Nature* **2003**, 421, 427-431.
5. Rothemund, P. W. K., Folding DNA to create nanoscale shapes and patterns *Nature* **2006**, 440, 297-302.
6. Qian, L. L.; Wang, Y.; Zhang, Z.; Zhao, J.; Pan, D.; Zhang, Y.; Liu, Q.; Fan, C. H.; Hu, J.; He, L., Analogic China map constructed by DNA *Chin. Sci. Bull.* **2006**, 51, 2973-2976.
7. Fujibayashi, K.; Hariadi, R.; Park, S. H.; Winfree, E.; Murata, S., Toward reliable algorithmic self-assembly of DNA tiles: A fixed-width cellular automaton pattern *Nano Lett.* **2008**, 8, 1791-1797.
8. Ke, Y. G.; Lindsay, S.; Chang, Y.; Liu, Y.; Yan, H., Self-assembled water-soluble nucleic acid probe tiles for label-free RNA hybridization assays *Science* **2008**, 319, 180-183.
9. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B., Addressing single molecules on DNA nanostructures *Angew. Chem.-Int. Edit.* **2007**, 46, 7956-7959.
10. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B., Programming multiple protein patterns on a single DNA nanostructure *J. Am. Chem. Soc.* **2008**, 130, 402-403.

11. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands *Nature* **1998**, 391, 468-471.
12. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 16768-16773.
13. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C. C.; Dervan, P. B., Completion of a programmable DNA-binding small molecule library *Tetrahedron* **2007**, 63, 6146-6151.
14. Puckett, J. W.; Muzikar, K. A.; Tietjen, J.; Warren, C. L.; Ansari, A. Z.; Dervan, P. B., Quantitative microarray profiling of DNA-binding molecules *J. Am. Chem. Soc.* **2007**, 129, 12310-12319.
15. Baird, E. E.; Dervan, P. B., Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids *J. Am. Chem. Soc.* **1996**, 118, 6141-6146.